#### DOCUMENT RESUME

ED 033 740 LI 001 809

TITLE Rules for Thesaurus Preparation.

INSTITUTION Cffice of Education, Washington, D.C.

Panel on Educational Terminology.

Pub Date Sep 69 Note 24r.

Available from Superintendent of Documents, U.S.

Government Printing Office, Washington,

D.C. 20402 (GPO FS.5.212; 12047,\$.20)

EDRS Price FERS Frice MF-\$0.25 HC Not Available from

EDRS.

Descriptors \*Guidelines, Indexing, Information

Networks, Information Retrieval,
\*Levicography \*Subject Index Terms.

\*Lexicography, \*Subject Index Terms,

\*Thesauri, \*Vocabulary

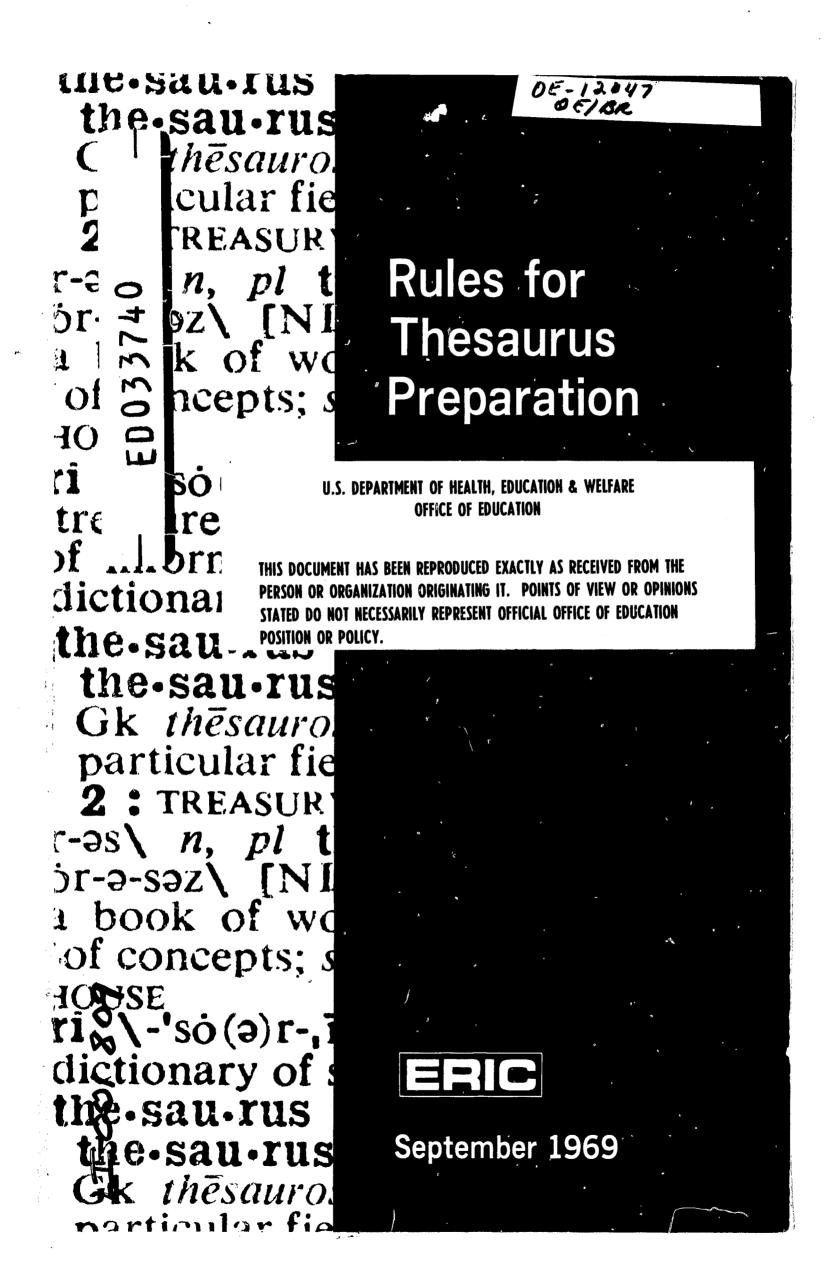
Identifiers Educational Resources Information Center,

ERIC, \*Thesaurus of ERIC Descriptors

### Abstract

ERIC

This is a revision of the "Rules for Thesaurus Preparation," published in October 1966. These rules are designed to help the Central ERIC staff and the staffs of the ERIC Clearinghouses make similar decisions related to the addition and modification of terms in the "Thesaurus of ERIC Descriptors," Second Edition. In addition to rules for expanding and modifying the Thesaurus, guidelines are set forth for developing subject-indexing vocabulary outside the scope of the Thesaurus, i.e., identifiers. Guidelines for the Descriptor Group Display and a description of the Rotated Descriptor Display are also included. (Author/JE)



Superintendent of Documents Catalog No. FS 5.212:12047 U.S. GOVERNMENT PRINTING OFFICE Washington: 1969

For sale by the Superintendent of Documents, U.S. Government Printing Office Washington, D.C., 20402 - Price 20 cents



This is a revision of the Rules for Thesaurus Preparation published in October 1966. In addition to rules for expanding and modifying the Thesaurus of ERIC Descriptors, Second Edition, guidelines are set forth for developing subject-indexing vocabulary outside the scope of the Thesaurus, i.e., identifiers. Guidelines for the Descriptor Group Display and a description of the Rotated Descriptor Display are also included.

Prepared and approved by the Office of Education Panel on Educational Terminology (PET)

James L. Eller, *Chairman* ERIC, U.S. Office of Education

Mary T. Fisher
North American Rockwell
Corp.

Terry L. Gillum System Development Corp.

Carter V. Good ERIC Consultant University of Cincinnati

Frederick L. Goodman ERIC Consultant University of Michigan Allen R. Lichtenberger National Center for Educational Statistics

Lawrence S. Papier ERIC, U.S. Office of Education

Mary J. Ruhl ERIC Clearinghouse on Exceptional Children

John A. Starkweather ERIC Consultant University of California

0018

# CONTENTS

	Foreword
1.0	Rules for Thesaurus Preparation
	Introduction
1.1	Descriptor Elements
1.1.1	$Descriptor\ Selection\ \_\_\_\_\_$
1.1.2	Cross-References
1.1.3	Ambiguity
1.2	Descriptor Construction
1.2.1	$Abbreviations \ \_\_\_\_\_$
1.2.2	Special Characters
1.2.3.	Word Form
1.2.4	$\it Numerals$
1.2.5	Descriptor $Length$
1.3	Descriptor Format
1.3.1	$Alphabetization \ \_\_\_\_\_$
1.3.2	$Descriptor\ Structures\ \_\_\_\_\_$
2.0	Descriptor Group Guidelines
	Introduction
2.1	Descriptor Group Display
2.2	Assigning New Descriptors to Descriptor Groups
3.0	Identifier Guidelines
	Introduction
3.1	Identifier Selection
3.1.1	Identifier Characteristics
3.2	Identifier Categories
3.3	References
4.0	Rotated Descriptor Display

### 1.0 Rules for Thesaurus Preparation

### Introduction

These rules are designed to help the Central ERIC staff and the staffs of the ERIC Clearinghouses make similar decisions related to the addition and modification of terms in the *Thesaurus of ERIC Descriptors*. Although these are not rules for indexing per se, they are invoked when an indexer or his supervisor finds that there is no term in the Thesaurus that is satisfactory for expressing an important concept found in a document. When this occurs, the indexer will record what he considers to be an appropriate term to express the concept. That term then becomes a candidate for inclusion in the Thesaurus. The rules are to be applied both in the selection of candidate terms and in the final decision to include the candidate term as a descriptor in the Thesaurus.

Rules should be regarded as useful in guiding those who ultimately must decide on the details of the Thesaurus. In this sense, a rule is analogous to a law in a "case law" system rather than in a "statutory law" system. Implementation of the rules is thus dependent on the development and analysis of specific cases involving them.

The Rules for Preparing and Up-dating Engineering Thesauri, prepared by the Engineers Joint Council, and the Manual for Building a Technical Thesaurus, prepared by Project LEX of the Office of Naval Research, have been especially useful in preparing this section.

### 1.1 Descriptor Elements

- 1.1.1 Descriptor Selection

  A descriptor is any single or multiword term which appears in the Thesaurus and which may be used for indexing a document. Rules for selecting descriptors are as follows:
- 1.1.1.1 Descriptors should represent important concepts found in the literature rather than concepts derived independently. They should also reflect the language used in the literature to describe such concepts.

- 1.1.1.2 Descriptors selected should have an agreed-upon meaning by relevant user groups and should be acceptable terminology for that user group. Acceptability will involve decisions as to obsolescence, negative connotations, colloquial usage, and other factors. (See also 1.1.3)
- 1.1.1.3 Since frequency of occurrence of terms is a factor in establishing descriptors, records should be kept of the number of times a candidate term has been used in indexing and/or searching.
- 1.1.1.4 Multiword descriptors (bound terms, precoordinated terms, and others) should be used whenever single-word descriptors cannot describe a concept adequately or provide effective retrieval. Many problems of this type can be solved by the careful application of rule 1.1.1.1 above. The following points should also be considered:
- 1.1.1.4.1 Use of a multiword descriptor is justified if any of the individual words in the multiword descriptor can combine so frequently with other descriptors as to produce many false coordinations.
- 1.1.1.4.2 Use of a multiword descriptor to represent a unique concept is justified if the individual words of that multiword term are also unique descriptors which, when coordinated with each other, represent concepts different from the one intended by the multiword term.

Example: STUDENTS
TEACHERS
STUDENT TEACHERS

1.1.1.4.3 If a single-word term (used as a substantive) is so general as to be virtually useless in searching (e.g., SCHOOLS), consider the use of that term with another term (e.g., SECONDARY SCHOOLS).

1.1.1.4.4 Multiword descriptors, like single-word descriptors, must be carefully considered for placement in descriptor hierarchies.

1.1.1.4.5 Do not use inverted entries. (See section 4.0)

Examples: DEVELOPMENT,

EMOTIONAL EDU-CATION, ADULT

are not valid descrip-

tors;

EMOTIONAL DE-VELOPMENT ADULT

**EDUCATION** 

are valid descriptors.

1.1.2 Cross-References

1.1.2.1 Use (USE)

The "USE" reference is utilized in two situations:

1.1.2.1.1 To indicate preferred usage

Example: Advanced Education

USE HIGHER EDUCATION

1.1.2.1.2 To cross-reference an abbreviation

Example: ETV

USE EDUCATIONAL

**TELEVISION** 

(*Note*: Inverted entries are not to be included in cross-referencing. For example, SCHOOL BUILDINGS would not be cross-referenced with Buildings, School.)

1.1.2.2 Used For (UF)

The mandatory reciprocal of the "USE" reference is the UF reference.

Examples: HIGHER EDUCATION

UF Advanced Education EDUCATIONAL TELE-

VISION UF ETV





1.1.2.3 Broader Term (BT) and Narrower Term (NT)

The broader term (BT) and narrower term (NT) cross references are employed to indicate any class relationships that may exist among descriptors. Where descriptors represent concepts that are included within the class of concepts represented by another descriptor, this relationship is shown by the broader term reference.

Examples: ORAL READING
BT READING
SILENT READING
BT READING

Each broader term reference requires a corresponding reciprocal narrower term reference:

Example: READING
NT ORAL READING
SILENT READING

(Note: A given descriptor may be a member of more than one class. If the descriptor is a member of more than one class, reciprocal references must be made to show these relationships to the next given level of the affected classes.)

Example: STUDENT TEACHERS
BT STUDENTS
TEACHERS

1.1.2.4 Related Term (RT)

The related term (RT) cross reference is employed as a guide from a given descriptor to other descriptors that are closely related conceptually but that do not possess class relationships as in 1.1.2.3 above. In general, the RT cross reference is given for the convenience of the user who, in examining one descriptor, needs to be reminded or informed of the existence of a related descriptor. A reciprocal relationship should always be listed.

### 1.1.3 Ambiguity

Parenthetical qualifiers and scope notes should be used when the intended usage of a descriptor has not been made explicit by cross references.

### 1.1.3.1 Parenthetical Qualifiers

A parenthetical qualifier identifies any particular indexable meaning of a homograph. This meaning is given the status of a descriptor by listing the homograph and its parenthetical qualifier in the thesaurus display. The homograph and its parenthetical qualifier are considered inseparable in indexing or searching. For any given homograph, there may exist as many descriptors consisting of homograph plus parenthetical qualifier as there are unique indexable meanings for that homograph. If a descriptor is judged to be a homograph, it must have a parenthetical qualifier.

Examples: GRADE (INCLINE)

GRADES

(SCHOLASTIC)

POSITION (LOCATION)
POSITION (TITLE)

Do not use another homograph as a parenthetical qualifier.

Example: do not use GRADES

(MARK)

but rather, GRADES (SCHOLASTIC)

One of the reasons for restricting the use of parenthetical qualifiers to homographs is to preclude the use of inverted entries.

#### 1.1.3.2 Scope Notes (SN)

A scope note is a brief statement appearing in the Thesaurus of the intended usage of a descriptor. It need not provide a formal definition. Scope notes are generally used in two situations:

1.1.3.2.1

To restrict the usage of a descriptor.

Example: OUT OF SCHOOL YOUTH

SN A child of compulsory school age who has been excused from attending school, or a child over 16 years of age who is out of school legally

1.1.3.2.2

To explain the usage of a descriptor which may have different meanings in several areas of education.

Example: ACCELERATION

SN The process of progressing through the school grades at a rate faster than that of the average child

Restricting usage as in 1.1.3.2.1 above will provide a useful retrieval term in instances where a questioner is interested only in experimental studies. A restricted usage might similarly be applied to "review" or "questionnaire study." Restricted usage will also give local ERIC units an opportunity to indicate precisely which additional usages must be provided for in a revised version of the Thesaurus. Explaining usage as in 1.1.3.2.2 above is quite arbitrary at this stage of thesaurus development, but will insure consistent use of a descriptor in indexing or searching.

### 1.2 Descriptor Construction

### 1.2.1 Abbreviations

In general, abbreviated terms should be avoided as descriptors since they may not be universally understood. Abbreviated and unabbreviated forms, such as acronyms, of a given term must be treated as synonyms and cross referenced. (See rules 1.1.2.1 and 1.1.2.2)

Special Characters

Special characters are any nonnumeric, nonalphabetic symbols appearing in the thesaurus display.

Special characters are permitted in scope notes within the limits of machine character availability. Right and left parentheses are the only special characters allowable in descriptors. (See rule

1.1.3.1)

1.2.3

Word Form In choosing between singular and plural noun forms, the precedent long established by major indexing and subject cataloging operations will be followed. Generally, a useful rule of thumb may be applied as follows: use the plural form when the proposed term is a "count noun," that is, a noun about which one should ask "how many?" (e.g., devices such as TEACHING MACHINES, PROJECTORS); use the singular form for "mass nouns," that is, those about which one should ask "how much?" (e.g., DELINQUENCY); use the singular for specific processes, properties, or conditions. Common usage should be followed for term types not covered in the above general rule (e.g., use DEMOCRACY not DEMOCRACIES).

1.2.4 Numerals
Numeric elements of descriptors will be displayed
as Arabic numerals and will follow the alphabetic
element of the descriptor.

Example: GRADE 4
not 4TH GRADE
or FOURTH GRADE

1.2.5 Descriptor Length
Descriptors are limited to 34 characters including spaces.

1.3	Descriptor Format
1.3.1	Alphabetization Alphabetize terms letter by letter.
1.3.1.1	Ignore all spaces between words.
1.3.1.2	Alphabetize according to the following sequence:
1.3.1.2.1	Left parenthesis
1.3.1.2.2	Numerals in usual order
1.3.1.2.3	Letters in usual order
1.3.2	Descriptor Structures Typical thesaurus term structures and required format:
	DescriptorACCELERATION
	Scope NoteSN The process of progressing through the school grades at a rate faster than that of the average child

Related Terms \_\_\_\_RT ACCELERATED
COURSES
ACCELERATED
PROGRAMS
ADVANCED
PLACEMENT
ADVANCED
STUDENTS
FLEXIBLE
PROGRESSION
TIME FACTORS
(LEARNING)

ERIC POPULATION FOR

8

Descriptor and parenthetical qualifier \_GRADES (SCHOLASTIC) Used for \_\_\_\_\_UF College Grades (Scholastic) **High School Grades** (Scholastic) Narrower Term \_\_\_NT GRADE POINT **AVERAGE** Broader Term \_\_\_\_BT ACHIEVEMENT RATING Related Terms \_\_\_\_RT ACADEMIC **ACHIEVEMENT CLASS AVERAGE CREDITS GRADING** REPORT CARDS SCORING Use Reference to an acceptable descriptor \_\_\_\_\_College Grades (Scholastic) **USE GRADES** (SCHOLASTIC) **High School Grades** (Scholastic)

**USE GRADES** 

(SCHOLASTIC)

## 2.0 Descriptor Group Guidelines

### Introduction

The Descriptor Group Display aids indexing, searching, and lexicographic analysis of descriptors. It permits browsing, shows descriptor relationships and usage, and helps to structure new descriptors. The Descriptor Group Display is designed as an aid for those using the Thesaurus and not as a classification scheme.

### 2.1 Descriptor Group Display

The Descriptor Group Display consists of three sections: (1) a list of the 52 Group headings with their respective three-digit codes, (2) a compilation of Descriptor Group Scope Notes which define each Descriptor Group, and (3) the Groups with descriptors. To provide an index or key to the Descriptor Groups, the three-digit group code appears two spaces to the right of each main descriptor in the Descriptor Listing.

### Example:

ADMINISTRATOR ATTITUDES 040 ATTITUDES 040

UF ADMINISTRATOR OPINION ATTITUDES ADMINISTRATION PERSONNEL ADMINISTRATOR CHARACTERISTICS ADMINISTRATOR CHARACTERISTICS ADMINISTRATOR EVALUATION EMPLOYERS

ADMINISTRATOR EVALUATION ENPLOYERS

ATTITUDES 040

ACADEMIC ASPIRATION ADMINISTRATOR ATTITUDES ADMINISTRATOR RESPONSIBILITY ANTI INTELLECTUALISM

# 2.2 Assigning New Descriptors to Descriptor Groups

A descriptor is assigned to only *one* Descriptor Group. Each new descriptor is assigned to a Descriptor Group by an indexer. The following guidelines are designed to assist the indexer in assigning descriptors to the proper Descriptor Group:

- 2.2.1 The assignment of a descriptor to a group should be considered in relation to the entire field of education rather than to a specialty.
- 2.2.2 Consult the existing Group Display to determine the disposition of similar, parallel, comparable, and analogous concepts.



ERIC

10

- 2.2.3 Consult the Descriptor Group assignments of terms cross-referenced from the new term.
- 2.2.4 Where terms appear to be logically related to more than one Group, choose the least abstract Group. (For example, READING ABILITY is assigned to READING rather than ABILITIES.)
- 2.2.5 Consider the word form of the descriptor since it may suggest a choice among Groups. (For example, COSMETOLOGY is assigned to OCCUPATIONS; COSMETOLOGISTS is assigned to PERSONNEL AND GROUPS.)

### 3.0 Identifier Guidelines

#### Introduction

The major purpose of identifiers is to provide terms of a specialized nature in addition to the descriptors in the *Thesaurus of ERIC Descriptors*. Unlike descriptors, identifiers are not cross-referenced or structured. They do not appear in the Thesaurus.

Some examples of identifiers: the name of a specific project (Project Head Start), geographical locations, persons, trade names, and educational or statistical tests. Identifiers refer to the subject of the indexed document.

### 3.1 Identifier Selection

- 3.1.1 Identifier Characteristics
- 3.1.1.1 Identifiers are not subject to lexicographic analysis.
- 3.1.1.2 Identifiers are neither structured nor cross referenced.
- 3.1.1.3 Identifiers are not entered into the structured vocabulary of the Thesaurus.
- 3.1.1.4 Parentheses, semicolons (field delimiters), and asterisks (major identifier indicators) are the only special characters permitted with the identifiers.
- 3.1.1.5 Identifiers are not abbreviated. (See 3.2.1.1 for exceptions)
- 3.1.1.6 Identifiers are limited to 50 characters including spaces. Avoid the use of over 50 spaces as identifiers are automatically truncated at the end of 50 spaces (e.g., Carnegie Institute Curriculum Study Center Curriculum in English should be entered as two separate identifiers: Carnegie Institute Curriculum Study Center; CICSC (Curriculum in English). The logical removal of prepositions, articles, and conjunctions to bring the identifier within the 50 space limitation is permitted.

3.1.1.7 Identifiers are entered with initial capital letters and lowercase (e.g., Department of Agriculture). Prepositions, articles, and connectives are in lowercase. Acronyms are in all uppercase.

### 3.2 Identifier Categories

### 3.2.1 Acronyms and Coined Terminology

3.2.1.1 Acronyms or initial letter abbreviations are entered in all uppercase in addition to the written-out version. Acronyms should not be used for major identifiers, the written-out version being used for any major designation.

AIMS

Access to Instructional Materials and Services

ARISE

Adult Referral Information Service in Education

CATE

Creative Application of Technology to Education

**CLASS** 

Cultural Language Activity Summer School

3.2.1.2 Coined terminology, or invented words, that are not broadly recognized or accepted by a subject specialty.

Pitcomings Shortfalls Failments Mitre Grammar

### 3.2.2 Organizations

The use of organizational names as identifiers is governed by the following rules:

3.2.2.1 If an organization is treated in a document from a subject standpoint, it is necessary to enter the organization as an identifier.

3.2.2.2 Where the organization is involved as both the subject of the document and the source of the document, the name should be used both as an identifier and as a source.

3.2.2.3 Specific types of organizations must be designated as follows:

3.2.2.3.1 Government departments at the Federal, State or local level when used as the subject of a document are entered:

Department of Agriculture
California Department of Labor
Department of Health Education and
Welfare

When a government department is indicated as an acronym, a double entry is made:

DHEW

Department of Health Education and Welfare

3.2.2.3.2 School districts used as the subject of a document are entered:

Tangipahoa Parish Louisiana School System

Long Beach California Unified School District

Polk County Florida Schools

For school districts designated by number; such as Kansas School District #259, convert the number to the proper geographic location:

Wichita Kansas School District

3.2.2.3.3 Educational organizations and names of educational centers are entered:

University of Iowa University of Pittsburgh Parent Teacher Association

When an educational organization is indicated as an acronym, a double entry is made:

ERIC

PTA
Parent Teacher Association
NEA
National Education Association
CAL
Center for Applied Linguistics

3.2.2.3.4 Community organizations are entered:

Cleveland Institute of Art

Los Angeles Chamber of Commerce

3.2.2.3.5 Industrial organizations are entered:
International Business Machines
Corporation
Westinghouse Corporation

3.2.2.3.6 Professional organizations are entered:

American Association on Mental

Deficiency

American Psychological Association

When a professional organization is indicated as an acronym, a double entry is made:

AAJC
American Association of Junior
Colleges
APA
American Psychological Association
MLA
Modern Language Association of
America

3.2.2.3.7 Foundations are entered:

Ford Foundation

Kaiser Foundation

Shell Company Foundation

3.2.3 Geographic Locations
Geographic locations are entered only if they reflect the subject content of the document.

Alberta Florida
Chicago Germany
Eau Claire County New York City Region
Englands Provinces South America
Europe Southwest

3.2.3.1 Where a city is not well enough known to stand by itself, enter both the name of the city and the State as two separate identifiers.

Memphis

Tennessee

Hollywood California

Hollywood Florida

3.2.3.1.1 Do not separate city and State with a comma; use a semicolon.

Memphis; Tennessee

3.2.4 Tests and Testing Programs

Tests and testing programs are entered:

Rorschach Index of Repression

When a test or testing program is entered as an acronym, a double entry is made:

ACT

American College Testing Program

**EPPS** 

Edwards Personal Preference Schedule

SCAT

School and College Aptitude Test

- 3.2.5 Textbooks, Trade Names, and Equipment
- 3.2.5.1 Textbooks and commercially published test series are entered:

Gregg Shorthand (Simplified Edition)
Uralic and Altaic Series

3.2.5.2 Trade names or proprietary designators are entered:

Miami Linguistic Readers Scott Foresman Basal Reader Learning to Read and Spell Reading in High Gear System for Success

3.2.5.3 Equipment names and numbers are entered:

Lisp 1 5

Autotutor Mark II

ERIC

3.2.6 Legislation and Assistance Programs

3.2.6.1 Legislation referred to by proper name is entered:

McAteer Act
Taft Hartley Act
Elementary and Secondary Education
Act

When legislation is indicated as an acronym, a double entry is made:

ESEA Title III
Elementary Secondary Education Act
Title III

3.2.6.2 Assistance programs that are designated by some authority at the Federal, State or local level are entered:

California Assembly Bill 1967

When assistance programs are indicated as acronyms, double entries are made:

ESEA Title I Program
Elementary Secondary Education Act
Title I Program

3.2.7 Other Identifiers

3.2.7.1 Specific names of methods and theories are entered:

Montessori Method Logogen System Chomsky Competency Model

3.2.7.2 Names of people used as the subject of a document are entered:

Ernest Hemingway Stephen Crane Plutarch Shakespeare J P Jones

3.2.7.2.1 Do not use a period following the authors' initials; use a space.

3.2.7.3 College or university course catalog listings are entered as follows:

English 2600

3.2.7.4 Indian tribes are entered:
Southern Paiutes
Navahos
Shawnees

3.2.7.5 Conferences, meetings, seminars, symposia, etc. are entered:

National Reading Conference Manpower Policy and Program Seminar

When conferences, etc., are entered as acronyms, double entries are made:

NDEA Language Institute
National Defense Education Act
Language Institute

3.2.7.6 Projects are entered with the word *Project* leading the entry:

Project Talent Project Head Start Project English

### 3.3 References

3.3.1 Government Departments
United States Government Organization Manual
1968-69. Annual rev. Washington, Government
Printing Office, 1968. 842 pp.

3.3.2 Educational Organizations
Education Directory 1966-67. Annual rev. Washington, Government Printing Office.

Part 1 State Governments. 1967. 158 pp. (OE-20004-67)
Part 2 Public School Systems. 1968. 209 pp. (OE-20005-68)
Part 3 Higher Education. 1967. 252 pp. (OE-50000-67)
Part 4 Education Associations. 1968. 140 pp. (OE-10001-67)

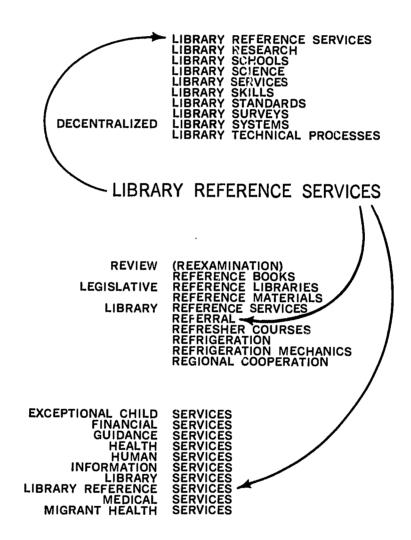
- 3.3.3 Geographic Locations
  Rand McNally Road Atlas—United States, Canada,
  Mexico. 44th annual ed. Chicago, Rand McNally,
  1968. 114 pp.
- 3.3.4 Tests and Measurement
  Buros, Oscar K., ed. Tests in Print: A comprehensive bibliography of tests for use in education, psychology, and industry. New Jersey, Gryphon, 1961.

Buros, Oscar K., ed. Mental Measurements Year-book. 6th ed. New Jersey, Gryphon, 1965.

### 4.0 Rotated Descriptor Display

The Rotated Descriptor Display is an alphabetical index to all words that form the descriptors in the Thesaurus. The word order within the descriptor is not affected. This display makes unnecessary inverted cross references in the Descriptor Listing.

An example is given below as it appears in the Rotated Descriptor Display.



### U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE OFFICE OF EDUCATION WASHINGTON, D.C. 20202

OFFICIAL BUSINESS



POSTAGE AND FEES PAID U.S. DEPARTMENT OF H.E.W.

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE Office of Education/Bureau of Research OE\_12047